

[12]发明专利申请公开说明书

[21]申请号 99119072.6

[43]公开日 2000年3月29日

[11]公开号 CN 1248864A

[22]申请日 1999.9.14 [21]申请号 99119072.6

[30]优先权

[32]1998.9.15 [33]FR [31]9811495

[71]申请人 汤姆森多媒体公司

地址 法国布洛里

[72]发明人 菲利普·吉约特

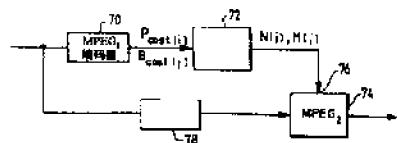
[74]专利代理机构 中科专利商标代理有限责任公司
代理人 朱进桂

权利要求书3页 说明书10页 附图页数2页

[54]发明名称 图像压缩方法以及实现这种方法的装置

[57]摘要

本发明涉及一种图像压缩方法，在方法中图像按照图组被编码，每一个图组包括 N 个图像，具有按帧内模式编码的 I 图像、作为内部图像 I 或前一幅图像 P 的函数预测的 P 图像，每个 P 图像前面或后面跟着 n 个双向预测 B 图像，n 可以为 0。数值 M=n+1 表示图组的结构。利用测试编码(70)确定至少一个表明要按照图组编码的源图像特征的参数，并且 M 和 N 依赖于这些参数。在测试编码过程中，已定义的值被赋予 N, M 以及量化区间 Q。



ISSN 1008-4274

权利要求书

1. 一种按照图组(GOP)编码图像的压缩方法图像，其中每个图组包含N个图像，N代表图组的长度，它包含以在帧内模式编码的I图像，在内部I图像或前一帧P基础上预测的P图像，每一个前面或后面有n个双向预测B图像，n有可能为0，等于n加一的数字M代表了图组的结构，其特征在于，至少有一个参数被确定来代表要被根据一图组编码的源图像的特征，并且这个图组的长度N和结构M取决于这个参数或这些参数。
2. 根据权利要求1的方法，其特征在于，描述源图像的特征的参数是通过在被编码数值被分配给N、M和量化区间Q的过程中利用测试编码而确定的。
3. 根据权利要求2的方法，其特征在于，测试编码在开环下完成。
4. 根据权利要求2或3的方法，其特征在于，为了描述源图像的特征，确定在测试编码中获得的描述P图像特征的参数(Pcost)和在测试编码中获得的描述B图像特征的参数(Bcost)。
5. 根据权利要求4的方法，其特征在于，数字N根据至少表明一个P图像的特征的参数而确定，以及数字M根据表明至少一个B图像的特征的参数而确定。
6. 根据权利要求4或5的方法，其特征在于，表明P和B图像特征的参数(Pcost, Bcost)是P和B图像编码的代价，例如平均代价。
7. 根据权利要求5和6的方法，其特征在于，在测试编码过程中，当对每一帧B图像编码的平均代价(Bcost)高于每一帧P图像编码的平均代价(Pcost)时，把值1赋给M，那么图组中不包含B图像。
8. 根据权利要求6或7的方法，其特征在于，在测试编码过程中，每一帧B图像编码的代价和对应的数字M被确定为与源图像的到达同步。
9. 根据权利要求8的方法，其特征在于，当在测试编码结束之前对于该组的大部分来说确定的M数字等于1时，则把数值1赋予该组的数字M。

10. 根据权利要求 8 或 9 的方法，其特征在于，当在测试编码结束之前对于该组的至少一个确定部分来说确定的 M 数字等于 1，并且当数字 M 对于前一个组等于 1 时，则把数值 1 赋予该组的数字 M。

11. 根据上述任何一项权利要求所述的方法，其特征在于，当在一组中发生场景改变时，该新的场景包括新组的 I 图像，如果场景的改变在受影响的组中出现，则缩短该受影响的组以在该新场景之前停止，距该场景开始处的距离至少等于 N 所许可的最小值，当在该受影响组中在场景改变之前的图像数目和在该受影响组之前的组的图像数目之和不大于 N 所许可的最大值时，受影响组的开始部分用于延长在其之前的组。

12. 根据权利要求 1 至 9 所述的方法，其特征在于，当在一组中发生场景改变时，该新的场景包括新组的 I 图像，受影响组和在其之前的组被重新排列，在改变后，使得它们每个组的长度接近于该组的和该组之前的组的长度的平均值。

13. 根据权利要求 6 所述的方法，其特征在于，通过根据 MPEG 类型标准，在 50 赫兹，使 N=12、M=3 以及 Q=15 执行的测试编码，数字 N 和 M 分别是对 B 和 P 图像编码的平均代价函数，其关系如：

$$(1) N = INT\left[\frac{389000 - P \text{cost}}{10000}\right] + 1, \text{ 当 } 12 \leq N \leq 30$$

$$(2) M = INT\left[\frac{179000 - B \text{cost}}{20000}\right] + 1, \text{ 当 } 1 \leq M \leq 7$$

INT 表示整数部分。

14. 根据权利要求 13 所述的方法，其特征在于， $1 < M < 7$ 。

15. 根据权利要求 13 或 14 所述的方法，其特征在于，当编码的代价 (Bcost) 大于 179000 时，数字 M 有如下关系来确定：

$$(3) M = 5 \cdot INT\left[\frac{P \text{cost}}{B \text{cost}} - 1\right], \text{ 当 } 1 \leq M \leq 7$$

16. 根据权利要求 1 至 6 中的任何一项所述的方法，其特征在于，数

字 M 在一组中变化。

17. 根据权利要求 2 至 6 中的任何一项所述的方法，其特征在于，该压缩是在测试编码之后执行的。

18. 根据权利要求 4 至 6 中的任何一项所述的方法，其特征在于，在已经确定表明第一 B 图像或第一 P 图像特征的参数之后，开始该压缩处理。

19. 根据权利要求 18 所述的方法，其特征在于，当编码图像的数目至少等于根据当前 P 图像所确定的数字 N 时，编码图像的形成被中断。

20. 根据权利要求 4、5、或 6 所述的方法，其特征在于，其包括用于执行测试编码并确定 M 和 N 参数以及从第一信道接收信息的编码信道，以便于适当地执行编码。

说 明 书

图像压缩方法以及实现这种方法的装置

本发明涉及一种图像压缩处理方法，也就是根据可变长图组对图像进行编码。

本发明特别涉及 MPEG 类型的一种方法，特别是 MPEG2 类型。虽然本发明并不仅限于这种标准，但在后面的描述中主要指 MPEG2 压缩方法。

压缩的原理在以下描述。

在视频 MPEG2 标准中，数字视频信号的压缩是对要编码图像的空间冗余度和时间冗余度来获得的。

空间冗余度主要由连续的三种运算得到的：一种通常叫做离散余弦变换并表示为 DCT (“离散余弦变换”) 的运算，一种对 DCT 变换得到的系数量化运算以及一种对 DCT 变换得到的量化系数进行可变长度编码(熵编码)的运算。

时间冗余度由运动补偿运算分析得到，补偿运算包含，对当前图像的每一块的变换，搜索在参考图像中最类似的块。对时间冗余度的分析可以决定了一场变换向量，一般被称为运动矢量。还有由当前图像信号和由运动补偿预测的图像信号的差别决定的预测误差。预测误差根据空间冗余度进行分析。

MPEG 编码是预测类型的编码，它使得有关的解码必须经常被重新初始化以保护信号不被传输错误或任何由解码器从一个程序转换到其他程序的信号中断。

为此目的，MPEG2 标准定期提供图像必须在空间模式中编码，这就是说仅仅按照一种利用空间冗余度的方式。图像在空间方式编码被称为帧内图像或 I 图像。

图像由时间冗余度决定编码有两种类型：一方面，图像由参考前一帧图像在前向预测基础上构造，另一方面，图像由参考前一帧和后一帧两帧

图像在前向预测和后向预测基础上构造。

由前向预测基础上构造的编码图像被称为预测图像或 P 图像，由前向预测和后向预测基础上构造的编码图像被称为双向图像或 B 图像。

一个 I 图像解码时而不需要参考其他图像。一个 P 图像解码时要参考前一帧 P 或 I 图像。一个 B 图像解码时依赖于前一帧 I 或 B 图像以及后一帧 I 或 P 图像。

I 图像的周期性定义了一组图像表示为 GOP (“图组”)。

在单个的 GOP 中，在 I 图像中包含数据的数量一般多于 P 图像中包含数据的数量，并且在 P 图像中包含数据的数量一般多于 B 图像中包含数据的数量。

在 50 赫兹时，GOP 表示为一帧 I 图像跟随一序列 B 图像和 P 图像，很多时候表示为以下序列：

I, B, B, P, B, B, P, B, P, B, B.

然而，标准并没有要求在 GOP 中提供 $N=12$ 帧图像，在一般情况下，也没有要求两帧 P 图像的距离 M 必须是 3。更精确地，距离 M 是在一帧 P 图像前面或后面的 B 图像的个数 n 加一，也就是， $M=n+1$ 。

数字 N 代表 GOP 的大小或长度，数字 M 代表它的结构。

本发明的结果来自于观察到可能在参数 M 和 N 上变化以增强压缩比和 / 或增强编码的质量。

由本发明的编码方法特征在于按照一个图组中至少有一个参数决定了要编码源图像的特征，并且这个图组的长度和结构的形成依赖于这个参数或这些参数。

在一个实施例中，描述源图像特征的参数在测试编码的帮助下确定的，它定义了值赋给 N, M 以及量化区间 Q。

测试编码被完成，例如，在开环情况下。

在一个特定的简单实施例中，描述 P 图像特征的测试编码中获得的一个参数 (Pcost) 以及描述 B 图像特征的测试编码中获得一个参数 (Bcost) 被分别确定，这些描述 P 图像和 B 图像特征的参数较好的指出了对 P 图像和 B 图像编码的平均代价。一幅图像的编码代价是指编码所必须的比特数 (包括头信息)。

在这种情况下，数字 N 可以取决于描述 P 图像特征的参数，数字 M 取决于描述 B 图像特征的参数。

在与本发明有关的实验中，对于不同类型图像的序列中，指出对于不同类型序列，一个最优的数字 N 存在，在对 P 图像最小的编码代价(或输入输出总和)下，并且一个最优的数字 M 存在，在对 B 图像最小的编码代价(或输入输出总和)下，这些代价在测试编码时获得。这些序列是由不同幅度，不同对象，不同空间定义以及不同内容的运动来区分的。

实验指出，此外，一个实用线形关系存在于最优数字 N 和 P 图像的输入输出总和。同样的，一个实用线形关系存在于最优数字 M 和 B 图像的输入输出总和。因此，知道了 P 和 B 图像的输入输出总和，很容易计算出能够得到最好的结果的数字 N 和 M。

在对应于 MPEG2 标准的一个例子中，50 赫兹，测试编码由 N=12，M=3 以及 Q=15 完成。N 和 P 图像的输入输出总和关系近似表示为下式：

$$(1) \quad N = INT\left[\frac{389000 - P \cos t}{10000}\right] + 1, \text{ 当 } 12 \leq N \leq 30$$

M 和 B 图像的输入输出总和关系近似表示为下式：

$$(2) \quad M = INT\left[\frac{179000 - B \cos t}{20000}\right] + 1, \text{ 当 } 1 \leq M \leq 7$$

也可以把 M 限制到 5。

在这些公式中，INT 代表整数部分。

限制 N 在 12 到 30 之间以及限制 M 最大值取 7 可得到编码器的一个简单实施例以及限制了程序变化的时间。在同样的目的下，也可能得到其他的约束或限制，特别是 M 在 GOP 中为常数并且/或它是 N 的约数。

在一个实施例中，如果 M 和 N 分别取值并且并不能同时满足约束，接近计算出来的并满足约束的值 M 和 N 被选择。在这种情况下，M 的值优先，也就是说如果要在许多组 M，N 中选择，将选择 M 的值最接近计算结果的组。

上述公式(2)满足 B_{cost} 不能超过 179800。在相反的情况下，也就是说如果 $B_{cost} > 179000$ ，实验指出有必要，在这个例子中， M 将由以下方式选择：

$$(3) M = 5 \cdot INT \left[\frac{P_{cost}}{B_{cost}} - 1 \right], \text{ 当 } 1 \leq M \leq 7$$

如果 B 图像的代价高于 P 图像的代价，选择 GOP 不包含 B 图像，这就是说 $M=1$ 。这是因为 P 图像表现出比 B 图像好的预测质量，并且在低代价的假设下，在这个例子中这样的 B 图像的存在将包含一个缺点。

由比特为单位每一帧 P 图像和 B 图像的代价被决定，例如，当这些图像出现时。在一个实施例中， M 和 N 的值由所有的测试编码的 P 和 B 图像的平均来选择，在测试编码 N 个源图像以后编码完全计算出来， N 由 P 图像的编码代价决定。在这种情况下，参数 M 可以在 GOP 中保持为常数。

在另一个实施例中，允许一个快速适应对场景内容的变化以及减少源图像到达和编码本身之间的延迟（这就允许一个小容量的缓冲存储器），编码开始当测试编码提供数据允许这个开始。于是，测试编码的第一个 B 图像提供了一个 M 数字使得编码开始并且数字 N 由测试图像的第一个 P 图像提供；在这种情况下，当 N 的值和 M 的值已知时编码开始。

在编码“在飞行中”的类型中，数字 M ，也就是说结构，可能在一个 GOP 中变化，这就允许对场景内容变化的快速适应。

在编码的逐步计算中，当在当前 GOP 已编码的图像数目至少等于衡量数字 N 时（由在以上例子中的 P_{cost} 来衡量），或在一个场景变化时 GOP 被中断。

为避免在图组之间参数的明显变化，可以证明值得做分离计算出来的值。例如，如果计算指出，对 GOP 大部分的长度，例如至少 80%，必须 $M=1$ ，然而，对于 GOP 的余下部分计算指出 M 应该大于 1， M 应该采用值 1，不管其他值，尽管计算指出必须有一个不同的值。

同样地，如果对前一个 GOP， $M=1$ ，并且如果对当前的 GOP，计算表明对于当前 GOP 大部分需要 $M=1$ ，例如至少 60%， M 应该接受值 1，尽管在以上公式(2)得到的结果，指出有一个不同的值。

已知当场景变化发生时，也就是说当视频图像序列发生不连续时，必须改变 GOP 图组在不连续的每一边使得新的开始于 I 图像的图组对应一个新的场景。

在一个实施例中，如果在一个图组中发生场景变化，新的场景建立一个新图组的 I 图像，如果场景变化发生在受影响的图组中，则受影响图组缩短使得在新场景之前从开始到至少等于 N 所允许的最小值的距离处停止。在前一个图组于是改变(变长或变短)，有可能要改变在这个 GOP 中先前计算出来的数字 M。

在一个变体中，经常优先选择在这种情况下受影响图组长度小于最小的允许 N，当场景改变发生在一个图组中，新的场景建立一个新图组的 I 图像，这个新图组的长度等于在它受影响的图组和前一个图组的长度的平均长度。在这个变化下，有必要改变为以前 GOP 校准的数字 M。

当两种改变都有可能时，例如当受影响图组长度小于 N 的最小允许值时，在这两种改变中一个选择可能由计算得出，对每一种改变，得到(M, N)对和改变之前的 M, N 对的距离并且选择距离最小的对。

为决定参数 N 和 M，要依靠参数的测量而不是输入输出总和的测量，例如，为决定 N，要使用 I 内部图像的能量。也可能决定运动或运动补偿误差的幅度，DFD(转移帧差)用来决定 M 和 N。

本发明的其他特征和优点将会在它的一些实施例描述中展现，这个描述在后面所附的图表，在其中

图 1 表示了 4.2.0 标准对应宏块。

图 2 是表示 DCT 变换的图表。

图 3 表示了一组图像，GOP，按照 MPEG 标准或类似标准。

图 4 至 7 是按照本发明方法的图表，和

图 8 是根据本发明方法实现的框图。

参考首先在图 1 到 3 目标是重申在 MPEG2 编码中用到的规则。

在 MPEG2 标准中，起点是一幅图像包含，在逐行方式，576 线每线 720 点。在隔行方式中，这幅图像包含两帧，每帧 288 线，同样每线 720 点。

每幅图像都被分成宏块，每个宏块由一个 16×16 的亮度方阵组成。每个宏块由 4 个 8×8 的亮度方阵构成。每 4 个亮度块与 2 个色度块有关(在

4.2.0 格式中), 每一个色度块是 8×8 点。其中一块表示色差即红色色度信号 C_r , 另一块表示色差即蓝色色度信号 C_b 。在 4.2.2 格式中, 每个亮度宏块对应 4 个 8×8 的色度块, 2 个块表示蓝色色度, 2 块表示红色色度。在 4.4.4 格式中, 每个亮度分量和色度分量包含有 4 个 8×8 块。

在图 1 所示的 4 个 8×8 的亮度块整体用标号 10 表示, 8×8 的色度块 12 和 14 分别代表蓝色和红色色度块。这是用 4.2.0 标准表示一个宏块。

每一个块用一种叫做离散余弦变换表示为 DCT 的变换编码, 它可把亮度块变换(例如)成代表空间频率的系数块。如图 2 所示, 源块 16 转换成一个 8×8 的系数块 18。块 18 左上角 20 对应空间零频率(块的均值), 从这个原点 20, 水平频率向右表示增加, 如箭头 22 表示, 同样, 垂直频率从上向下增加, 如箭头 22 表示。

对每一个宏块, 必须要选择编码的类型: 或者“帧内编码”或者“帧间编码”。帧内编码指对图像的源块进行 DCT 变换; 帧间编码指对表示源块与预测块差异的一个块, 或者是上一幅图像或下一幅图像的预测块进行 DCT 变换。

选择是部分地依赖于宏块所属的图像类型。图像有三种类型: 第一种是所谓的 I 或称内部类型, 对所有宏块都用帧内编码。

第二种类型是 P 或称预测类型。在这种图像中, 对宏块的编码可用帧内编码也可用帧间编码。如果对 P 型图像使用帧间编码, 是对这幅图像的当前块与前面的 I 型或 P 型图像产生的预测块的差异进行 DCT 变换。

第三种图像称之为 B 或双向类型。对宏块也可用帧间编码或帧内编码。帧间编码也是对当前块与预测块的差异进行 DCT 变换。这个预测块可由前向图像, 或后向图像, 或两者均可(双向的预测)产生。这就有可能对所称的前向或后向预测图像是 I 或 P 图像之一。

图 3 所示的一序列图像组成一组称为图组(GOP), 它包含 12 幅图像。第 1 个为 I 图像, 后面 11 个为 B 和 P 图像, 按如下顺序: B, B, P, B, B, P, B, B, P, B, B。

一个 GOP 由长度表示其特征, 这就是说一序列图像 N, 例如, 可在 12 到 30 之间, 并且结构参数 M 表示了两个 P 图像之间的距离, 也就是说跟随在一帧 P 图像后面 B 图像的个数加一。在这个例子中, 参数 M=3。同样

300-300 lb.

在例子中，这个数字 M 可能在 1 (没有 B 图像) 到 7 之间。而且规定了数字 M 必须是 N 的约数以简化解码器。

到目前为止，在编码器中图像编码时保持 N 和 M 为常数。

本发明的结果来自于观察到对于不同的编码图像序列存在最优的参数 M 和 N。这是因为，依赖于是否图像序列清晰度的大小，以及运动的程度，M 和 N 的最优值会明显的不同。用最优值在相同的质量下，需要最少的比特数。

紧接着，实验研究得到了在本发明指出对于对应于最小值 $Pcost$ 的已定义的图像序列 GOP 的最优 N_{opt} ，在这个序列下，有必要对 P 图像进行编码（包括头信息）。这个性质由图 4 所示，在横坐标为数字 N ，在纵坐标为对序列表示为 I 的值 $Pcost$ 。值 $Pcost$ 是编码 P 图像的比特数在序列 I 的平均值。可以看出，值 $Pcost(i)$ 由曲线 32 表示有一最小值 34，这时 N 的值是最优的 (N_{opt})。

同样的，对于最小值 $Bcost(i)$ 的已定义的图像序列 B 图像平均比特数的最优值 M ，表示为 i 。因此，在图 5 中，在横坐标为数字 M ，在纵坐标为对序列表示为 $Bcost(I)$ 。可以在图中看出，曲线 36 有对应 $M(M_{opt})$ 的最优值的最小值 38。

测量在 MPEG 通用的序列中进行，它们是“Horse”，“Flower garden”和“Mobcal”。“Horse”序列对应快速运动和良好清晰度，“Flower garden”序列对应良好清晰度和中速运动，“Mobcal”序列对应慢速运动和高清晰度。其他序列也被测试，如有快速运动和差的清晰度的皮船序列，一个篮球序列以及一个有匀速运动和良好清晰度的序列。

这里指出, 如果测试编码的图组有已定义的值 M , N 和量化区间 Q , 这些值没有必要对应所讨论的序列 I 的最优值, P 帧编码平均代价 $Pcost$ 和 B 帧编码平均代价 $Bcost$ 分别代表 N 和 M 。此外, 如图 5 所示, 一个简单的关系存在于对应序列 I 的数字 $Nopt$ 和给定 M , N 和 Q 的编码代价 $Pcost$ 之间。这个关系是线形的或近似线形的, 可以表示为直线 40(图 6), 在上面不同的点 42, 44 对应不同的序列。

在图 7 中在横坐标为 N_{opt} , 在纵坐标为编码代价 $Pcost$ (对于给定 N , M 和 Q)。每一个的点 52、54、56、等等对应一个给定的序列, 可以看出

300.000.00

这些点在一条直线 60 上。因此一个线形关系存在于 N_{opt} 和测试编码代价之间。

在测试编码时 M, N 和 Q 取以下的值:

M=12,

N=3 以及

Q=15。

N 和 M 的值满足以下关系式:

$$(1) \quad N = INT\left[\frac{389000 - P_{cost}}{10000}\right] + 1, \text{ 当 } 12 \leq N \leq 30$$

$$(2) \quad M = INT\left[\frac{179000 - B_{cost}}{20000}\right] + 1, \text{ 当 } 1 \leq M \leq 7$$

尽管, 从上述公式(2), 表明 M 应该在 1 到 7 之间, 从图 7 可以看出 M 可能被限制到 5。

实施本发明的框图在图 8 中表示。它包含了一个第一 MPEG2 编码器 70 目的是计算出测试编码或“第一遍通过”。测试编码由以上指出的恒定参数建立, 即, 在这个例子中: N=12, M=3 和 Q=15。这个测试编码器在这个例子中, 在开环下工作, 也就是说不用校准。

编码器 70 提供了 B_{cost} 和 P_{cost} 的值到一个转换器 72, 它计算出 P_{cost} 转换为 N_{opt} 以及 B_{cost} 转换为 M_{opt} , 正如在图 6 和图 7 表示的以及按照上述关系(1)和(2)。

这些 N 和 M 的值对一个图组计算, 如上面所述, 然后提供给 MPEG2 编码器 74 的控制输入端 76。

编码器 74 的输入数据与测试编码器 70 的输入是一样的。因此考虑到测试编码器 70 和转换器 72 的处理时间, 一个缓冲存储器 78 被提供, 这个存储器 78 在处理过程中保存数据。

在转换器 72 中, 必须检查从(1)和(2)中得到的 N, M 对是否与限制条件一致, 特别是 M 必须是 N 的约数。如果从计算得到结果的值不相符, N 和 M 的值将接受接近那些计算得到的值, 尤其是 M 的值。

转换器 72 也应考虑以下情况:

在第一种情况下，比较 Bcost 和 Pcost 的值，如果 Bcost 比 Pcost 要高，值 1 被赋给 M，GOP 中没有 B 图像。这是因为，在这种假设下，B 图像承受了一个高于 P 图像的编码代价；只保持 P 图像有 1 一个更高的预测质量。

在第二种情况下，转换器比较 Bcost 的值和 179000 并且，如果 Bcost 超过 179000，关系式(2)由以下关系式代替：

$$(3) M = 5 \cdot \text{INT} \left[\frac{P \cos t}{B \cos t} - 1 \right], \text{ 当 } 1 \leq M \leq 7$$

转换器 72 使得考虑两种与关系式(2)不同的特殊情况成为可能为取得一致的图像质量。

第一种情况如下：测试编码允许 M 的值至少为 2，但是，这个测试编码同样表现出在图组的大部分由 M 获得的中间值等于 1，例如至少 80%。在这种情况下，转换器 72 规定 M 等于 1。

第二种情况与第一种情况类似：测试编码允许 M 的值至少为 2，但是，这个测试编码同样表现出在图组的较少一部分由 M 获得的中间值等于 1，例如至少 60%（这个限制低于第一种情况），并且前一个图组是 M=1。在这种情况下，值 1 赋给 M。

对于这两种特殊情况，值 1 赋给了 M，结果来自于与本发明有关的实验计算，使得这些条件对于相同类型的序列在连续的图组上可获得具有良好一致性的质量。.

最后，转换器 72 考虑到在编码器中检测到的场景的变化或“截断”。当这样场景变化发生时，GOP 从一个新的场景开始，这就是说，当新的场景出现时，它归结于一个 I 图。

此外，由本发明的方法，当场景变化被检测到，前一个 GOP 和当前的 GOP 在以下的考虑基础上设定：

如果场景变化发生在一个 GOP 中在 12 帧图像后，当场景变化时新的 GOP 开始，前一个 GOP 于是被缩短。

如果，相反的，场景变化发生在一个 GOP 中在 12 帧图像前，就不可能限制前一个 GOP 所以在场景变化前终止，因为，在这种情况下，图像的

数目要小于规定的最小数字。前一个 GOP 和当前的 GOP 用以下方法改变，两种情况被区分。

在第一种情况下，场景变化发生的时刻是在场景改变之前，前一个 GOP 的帧数与当前 GOP 的帧数的和至少等于 30 的时刻。在这种情况下，前一个 GOP 被加长。

在第二种情况下，前一个 GOP 的帧数与当前 GOP 的帧数的和在场景改变之前大于 30。前一个 GOP 和当前 GOP 由这两个 GOP 的平均重新排列。

例如，如果前一个 GOP 为 $N=25$ 和 $M=2$ 以及场景变化发生在当前计算表明 $N=20$ 和 $M=3$ 的 GOP 的第八帧图像，由当前缩短的 GOP 加长的前一个 GOP 包含 33 个图像。当这个值超过了最大允许值(30)时，一个“平均值”被找到对应于两个 GOP 并且总的图像数目为 33，每一个 GOP 遵循强制限制。在这种情况下，可以看出选择可以在 $N=16$ 以及 $M=2$ 对前一个 GOP 和 $N=15$ 以及 $M=3$ 对场景变化前的 GOP 之间。长度 18 和 15 是靠近平均(16, 5)前一个图组(25)的长度以及当前受影响的图组长度(8)。

实验在场景、闪光和持续长度变化的 12 个不同序列进行，结果由传统的编码方法获得，对应恒定的 M 和 N 的值对比由本发明的方法变化在序列中 M 和 N 的值的结果。这些测试由计算得到一些结果。由 PSNR 参数(信噪比)来度量，在质量上增加 0.2dB 到 1.14dB。这个在 PSNR 上的增加对应着比特数节约为 2 到 22% 之间。

按本发明的方法可以用于任何类型的具有 I、P 和 B 图像的视频图像压缩方法。它适用于记录，实时或离线，以及传输。

本方法不仅仅限于在编码前 GOP 的大小已固定的情况。它还适用于当参数 M 和 N 对每一帧图像计算，编码代价实时计算出。在这种情况下， M 会在一个 GOP 中变化，一个新的 GOP 开始，例如，在当前 GOP 中编码图像数目至少等于计算值 N 。 M 可能作为在 GOP 中图像复杂度的函数而变化。

在这种情况下，没有必要在缓冲存储器 78 中(它的容量可以被减少)存储全部的 GOP， M 和 N 的限制也被减少，由 MPEG2 标准单独指出；利用场景变化的限制也不很严格。

99.09.14

说 明 书 附 图

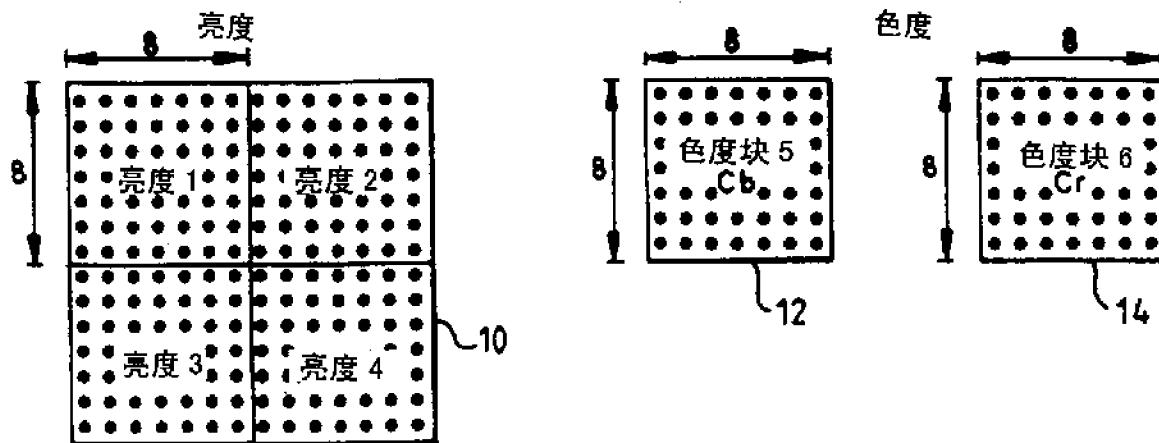


图 1

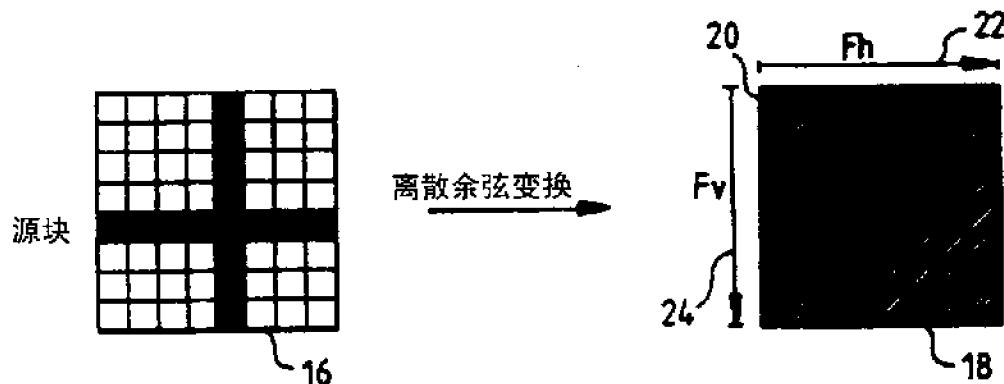


图 2

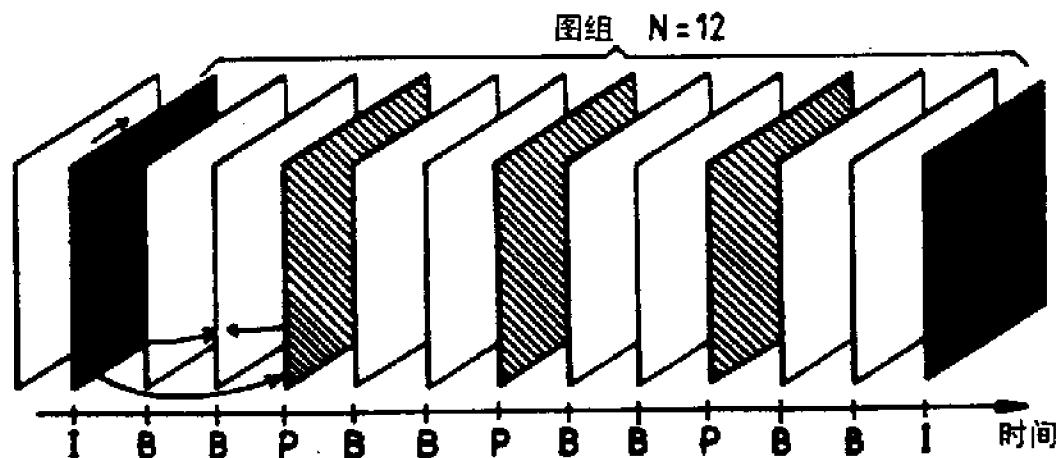


图 3

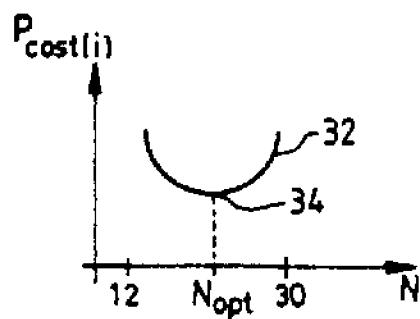


图 4

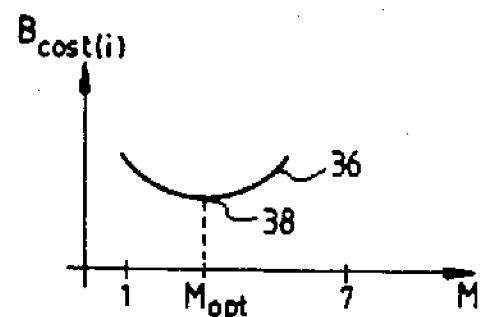


图 5

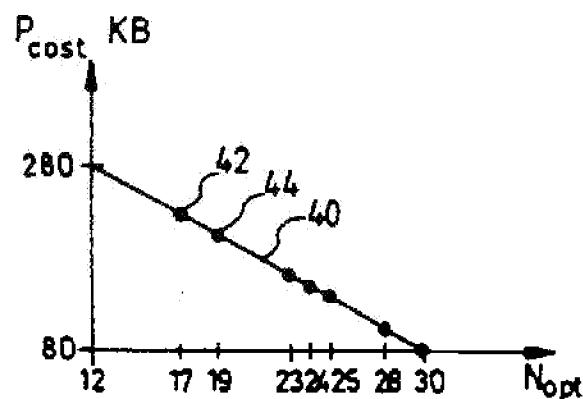


图 6

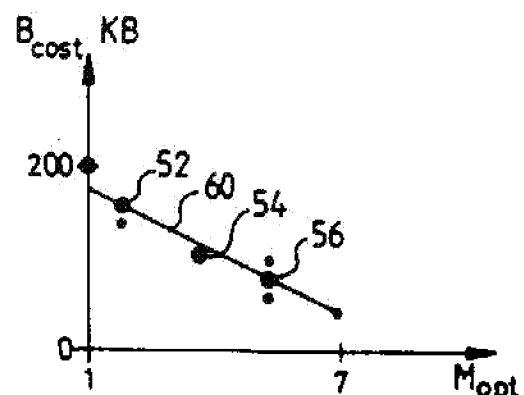


图 7

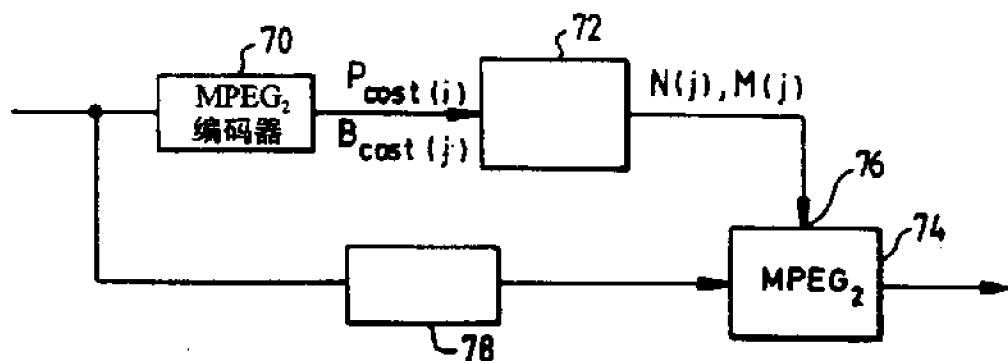


图 8

02 - BR 9904108
04 - CN 1248864
07 - EP 987903
10 - FR 2783388
15 - JP 2000102021
20 - KR 2000023133
27 - ZA 9905802

The method involves coding images according to groups each of which comprises N images, with an I image coded in intra mode, P images predicted as a function of the intra image I or of the preceding P image. Each P image is preceded and followed by n bidirectionally predicted B images, n possibly being zero. The number M = n + 1 represents the structure of the group.

At least one parameter characterising the source images which are to be coded according to a group is determined using a test coding (70). The numbers N and M are made to depend on this parameter. In the course of the test coding, defined values are conferred on N, M and on the quantisation interval Q.